$$x_k = A_k x_{k-1} + B_k u_k + w_{k-1}$$
$$z_k = H_k x_k + v_k$$

Prediction $\hat{x}_k$

$$\hat{x}_k = A\widehat{x_{k-1}} + Bu_k$$

Optimal Estimation $\hat{x}_k^*$

$$K = \frac{Predicted\ Error}{Predicted\ Error + Measurement\ Error}$$
$$\boldsymbol{\hat{x}_k^* = \hat{x}_k + K(z_k - H\hat{x}_k)}$$
$$\hat{x}_k^* = \hat{x}_k + K(Hx_k + v_k - H\hat{x}_k) = \hat{x}_k + KHx_k + Kv_k - KH\widehat{x_k}$$
$$\hat{x}_k^* - x_t = \widehat{x_k} - x_t + KHx_k + Kv_k - KH\widehat{x_k}$$
$$e_k = x_t - \hat{x}_k^*$$
$$\widehat{e_k} = x_t - \widehat{x_k}$$
$$e_k = (I - KH)\widehat{e_k} - Kv_k$$

Covariance Matrix

$$P_k = Cov\big((I - KH)\widehat{e_k} - Kv_k\big) = (I - KH)\widehat{P_k}(I - KH)' + KRK^T$$

$$P_k = \widehat{P_k} - KH\widehat{P_k} - \widehat{P_k}H'K' + K\big(H\widehat{P_k}H' + R\big)K'$$

$$\frac{\partial P_k}{\partial K} = -2\big(\widehat{P_k}H'\big) + 2K\big(H\widehat{P_k}H' + R\big) = 0$$

$$\boldsymbol{K = \widehat{P_k}H'\big(H\widehat{P_k}H' + R\big)^{-1}}$$

Then

$$P_k = \widehat{P_k} - KH\widehat{P_k} - \widehat{P_k}H'K' + K\big(H\widehat{P_k}H' + R\big)K' = \boldsymbol{\widehat{P_k} - KH\widehat{P_k}}$$

Estimate Covariance Matrix

$$\widehat{e_k} = A_k x_{k-1} + B_k u_k + w_{k-1} - (A_k\widehat{x_{k-1}} + B_k u_k)$$
$$\widehat{e_k} = Ae_k + w_k$$
$$\widehat{P_{k+1}} = E[(Ae_k)(Ae_k)'] + E(w_k w_k')$$
$$\boldsymbol{\widehat{P_{k+1}} = AP_kA' + Q}$$